Time-delay quasars: scales and orders of magnitudes

Prasenjit Saha[1,2]

[1] Astronomy Unit, Queen Mary and Westfield College, University of London, London E1 4NS, UK
e-mail: `p.saha@qmul.ac.uk`

[2] Observatoire astronomique, 11, rue de l'Université, Strasbourg F-67000, France

Abstract. We can think of a lensed quasar as taking the Hubble time, shrinking it by $\sim 10^{-11}$, and then presenting the result to us as a time delay; the shrinking factor is of the order of fractional sky-area that the lens occupies. This cute fact is a straightforward consequence of lensing theory, and enables a simple rescaling of time delays. Observed time delays have a 40-fold range, but after rescaling the range reduces to 5-fold. The latter range depends on details of the lens and lensing configuration—for example, quads have systematically shorter rescaled time delays than doubles—and is as expected from a simple model. The hypothesis that observed time-delay lenses all come from a generalized-isothermal family can be ruled out. But there is no indication of drastically different populations either.

Key words. Gravitational lensing – quasars: general –

1. Introduction

Most of the observables in gravitational lensing (image positions and magnifications) are intrinsically dimensionless. The exception is the time delay between images, which takes its dimensionality straight from the universe:[1] $\Delta t \propto H_0^{-1}$. This remarkable fact is the essential reason for much research effort going into measuring time delays. The observations have been increasingly successful—in 1995 there was but one controversial time delay, currently there are nine non-controversial ones. These are summarized in Table 1 below.

But curiously, even as the image and time delay data have improved, the error bars on the inferred H_0 have not. As an example, consider 0957+561. Between Kundić et al. (1997) and Oscoz et al. (2001) the time-delay value changed by only 2%. But meanwhile, whereas Kundić et al. (1997) quote $H_0 = 64 \pm 13$ (95% confidence) in the usual units of $\mathrm{km\,s^{-1}\,Mpc^{-1}}$, Bernstein & Fischer (1999) with more imaging and more modelling conclude that the data imply only 77^{+29}_{-24}, while Keeton et al. (2000) assert that further data on the lensed host galaxy invalidates all previously published models, and they decline to give an H_0 estimate at all. Basically, the problem is that simple lens models are unable to fit the images to the mas-level demanded by current data, while more complicated models can fit the data but are non-unique and can produce identical observables from very different values of H_0.

Modellers have responded to this dilemma with two strategies. One is to try to identify simple models that both have enough parameters to fit or nearly fit the data and can

be justified on galactic-structure grounds; Kochanek (2003) is typical of these. The other strategy is to try to explore the space of all plausible models allowed by the data; Raychaudhury et al. (2003) is a recent example. For a review by authors representing different points of view see Courbin et al. (2003).

In the current context of good data and active modelling but no consensus on models, it is interesting to step back and pose some questions that tend to get obscured in the details of modelling. First, we can think of the purpose of modelling time-delay lenses as being to discover one dimensionless number, the factor relating Δt and H_0^{-1}. What contributions to this number are well-constrained and what are poorly constrained? What range of values do the data imply for the poorly-constrained part? Is that range systematically different for doubles and quads, and/or for isolated lensing galaxies versus interacting galaxies? And is that range consistent with what we expect from popular models? Nine systems is a small sample, but it is enough to provide preliminary answers to these questions, and to do so is the aim of this paper.

2. A scaling relation for time delays

In lensing theory the arrival time can be written as

$$t(\vec{\theta}) = (1 + z_\mathrm{L})\frac{D_\mathrm{L}D_\mathrm{S}}{cD_\mathrm{LS}} \left[\tfrac{1}{2}|\vec{\theta} - \vec{\beta}|^2 - \psi(\vec{\theta}) \right] \tag{1}$$

where the symbols have their usual meanings. For convenience, let us abbreviate this expression. First, we write $\tau(\vec{\theta})$ for the expression inside square brackets. The factor outside square brackets equals H_0^{-1} times a dimensionless distance factor D

---

[1] This point appears to have been first emphasized by Nityananda (1990), although it is implicit already in Refsdal (1964).

(say) that depends on redshifts and (weakly) on cosmology, but not on H_0; for small z_L and large z_S, $D \simeq z_L(1 + z_L)$. We note further that only differences in arrival-time between images are observable. Hence observable time delays have the form

$$\Delta t = H_0^{-1} D \Delta\tau \tag{2}$$

We expect that $\Delta\tau$ will be of the same order as $|\vec{\theta} - \vec{\beta}|^2$ but a few times smaller, the precise value depending on details of lens and lens configuration. For an observed lens we might predict

$$\Delta\tau \sim (\theta_1 + \theta_2)^2 \tag{3}$$

where θ_1, θ_2 are the θ values of the first and last images to arrive. To focus attention on the proportionality factor, I propose to consider the dimensionless quantity

$$\varphi \equiv \frac{\Delta\tau}{\frac{1}{16}(\theta_1 + \theta_2)^2}. \tag{4}$$

We can calculate φ from a lens model, but not directly from observations. We can, however, measure a related quantity, a scaled time delay

$$\Delta T \equiv \frac{\Delta t}{\frac{1}{16}(\theta_1 + \theta_2)^2 D} \tag{5}$$

directly from observations, and substituting equations (2) and (4) we see that

$$\Delta T = \varphi H_0^{-1}. \tag{6}$$

The factor $\frac{1}{16}$ is ad hoc, but it allows the following interpretation. Recall that the image separation in a galaxy lens is about twice the Einstein radius:

$$\theta_1 + \theta_2 \simeq 2\theta_E. \tag{7}$$

For an isothermal, the relation (7) is exact. But even for non-circular lenses, where θ_E is not strictly defined, the image configuration can be used to define an effective θ_E. Using (7) the denominator in (4) is $\pi\theta_E^2/(4\pi)$, i.e., the area of the Einstein ring as a fraction of the sky. In other words, if we scale the observed time delay by the lens's covering factor on the sky we get H_0^{-1} times a 'fudge factor' of the order of unity.

For isothermal lenses, φ ranges from 0 to 8, averaging $\frac{16}{3}$. To see this, recall that for isothermals, $\Delta\tau = 2\theta_E\beta$ and note that β could be anywhere in the Einstein ring. Hence $\langle\beta\rangle = \frac{2}{3}\theta_E$ and using (7) gives

$$\langle\Delta\tau\rangle_{\text{iso}} = \frac{1}{3}(\theta_1 + \theta_2)^2. \tag{8}$$

Equation (8) is interesting for comparison with non-isothermals, but for isothermals themselves, we can do better. Combining $\Delta\tau = 2\theta_E\beta$ with $\theta_1 - \theta_2 = 2\beta$, which isothermals also satisfy, allows us to define

$$\Delta T_{\text{iso}} \equiv \frac{\Delta t}{\frac{1}{2}(\theta_1^2 - \theta_2^2) D} \tag{9}$$

which equals H_0^{-1}.

Witt et al. (2000, hereafter WMK) show that $\Delta T_{\text{iso}} = H_0^{-1}$ is not restricted to isothermals but is valid for a large family of generalized-isothermal lenses, and argue that it will be generally applicable in nature. If so, φ could be eliminated altogether. We can readily test if this is the case.

Table 1. Summary of time-delay data.

Object	θ_1	θ_2	D	Δt
0957+561	5.23	1.03	0.49	423 ± 1[a]
0911+055	2.24	0.82	1.12	146 ± 8[b]
1520+530	1.21	0.39	1.21	130 ± 3[c]
2149–275	1.37	0.33	0.67	103 ± 12[d]
1608+656	1.53	0.58	1.20	77 ± 3[e]
1600+434	1.02	0.39	0.59	51 ± 4[f]
1830–211	0.68	0.30	1.47	26^{+5}_{-4}[g]
1115+080	1.37	0.95	0.39	25 ± 4[h,i]
0218+357	0.22	0.14	2.42	10 ± 1[j,k]

[a]Oscoz et al. (2001)   [b]Hjorth et al. (2002)   [c]Burud et al. (2002b)
[d]Burud et al. (2002a)   [e]Fassnacht et al. (2002)
[f]Burud et al. (2000)   [g]Lovell et al. (1998)
[h]Schechter et al. (1997)   [i]Barkana (1997)   [j]Biggs et al. (1999)
[k]Cohen et al. (2000)

3. Scaling the data

We now present the obvious comparison of the scaled time delays $\Delta T = \varphi H_0^{-1}$ with current data.

Table 1 lists the relevant quantities for the various time-delay systems. The time delays references are given in the table, and the other data are taken from the CASTLES survey and compilation by Kochanek et al. (1998). For quads, only the first and last images (that is, the longest time delay) are considered, to enable a simple comparison with doubles. There are some caveats to the values of θ_1 and θ_2: for 1830 and 0218 the lens-centre is very uncertain and hence θ_1, θ_2 are especially uncertain, for 1608 the lens is apparently an interacting pair of galaxies, and 0957 and 0911 are in clusters and hence have large lensing contributions from other galaxies.

Figure 1 shows ΔT against Δt for the currently known time-delay systems. Since error bars on time delays are typically a few percent they are not shown here. We notice three things:

– Whereas Δt ranges over a factor of 40, ΔT ranges over a factor of 5.
– No correlation is evident between ΔT and Δt. According to the shuffling test described in Appendix A, the trend is significant at the 75% level—i.e., not significant. (Meanwhile, Figure 2 shows how Figure 1 changes if we ignore all redshift information and simply set $D = 1$. The scatter increases, but again there is no significant trend.)
– If we assume that H_0^{-1} is ~ 15 Gyr, then the range of φ is 1.5–2 for quads and about 2–6 for doubles. [R. Ibata (personal communication) drew attention to this separation from an early version of Figure 1.]

The various caveats above do not appear to affect these points.

We can also compare $\Delta T_{\text{iso}} = H_0^{-1}$ against the data to test whether the lenses belong to the generalized isothermal family studied by WMK. Figure 3 shows ΔT_{iso} against Δt for the same systems. We notice the following

– ΔT_{iso} (expected to be constant, since there is no φ factor) ranges over a factor of 5.
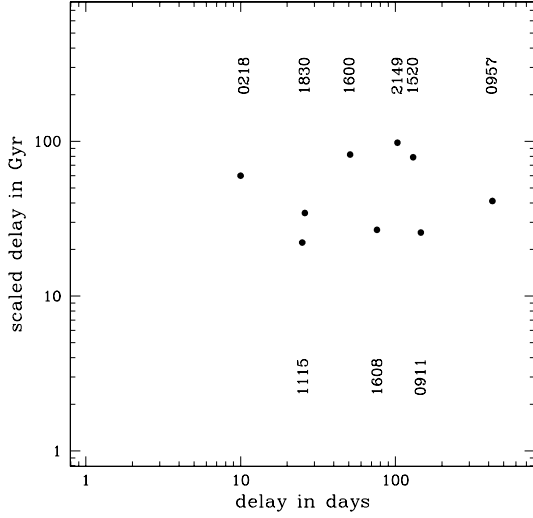
Fig. 1. Plot of the scaled time delay ΔT defined in Equation (5) against the observed time delay. The various lenses are labelled by their short names: quads are labelled below, doubles above.

Fig. 2. As in Figure 1, but omitting the D factor in the scaled time delay.

– Larger lenses tend to give lower ΔT_{iso}, and according to the shuffling test, this non-physical trend is significant at the 95% level.

– Of the nine lenses, only 1115, 1520 and 1830 even give $10\,\mathrm{Gyr} < \Delta T_{\mathrm{iso}} < 20\,\mathrm{Gyr}$, let alone a consistent $\Delta T_{\mathrm{iso}} = H_0^{-1}$.

Again we must keep in mind the caveats above, and also that the large external shear in 1115, 0957 and 0911 means that for these lenses ΔT_{iso} properly speaking requires a modification given in WMK but disregarded here. On the other hand, it seems unlikely that these caveats will solve the serious discrepancies we see. It appears more likely that most real lenses do not belong to the generalized isothermal family.

Whereas ΔT_{iso} is rejected, are other scalings possible that improve upon ΔT? L.L.R. Williams (personal communication) points out that the definition (Equation 5) of ΔT considers the size of the lens but not its asymmetry, and that if we multiply $(\theta_1 + \theta_2)^2$ in the definition by a further factor of $\sqrt{(\theta_1 - \theta_2)/(\theta_1 + \theta_2)}$ as a measure of asymmetry, then the scaled time delays would range over a factor of only 2.5, with no significant trend. But the meaning of such an asymmetry correction in terms of lensing theory is not known.

4. Modelling the range of φ

From the above, it appears that the scatter in φ reflects a range of mass profiles and source positions, and that its value must be inferred for each lens by detailed modelling. But without going into detailed models for nine lenses, we can at least check whether the observed range of φ is plausible.

Figure 4 shows such a check. The main plot is of φ against the area $(\theta_1 + \theta_2)^2$ for an example model (an elliptical isothermal potential plus external shear.) The value of φ is shown for different source positions, the two loops corresponding to source positions along the two caustics (actually just inside the caus-
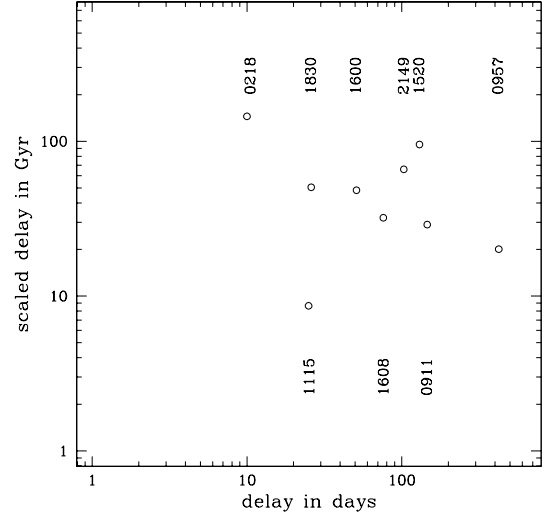

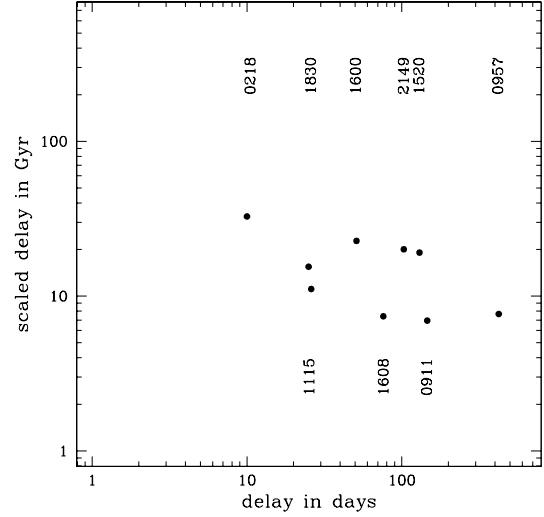**Fig. 3.** T_{iso} as defined in Equation (9) against the observed time delay. The non-physical trend is significant (see text), and hence the generalized isothermal models are rejected.

tics, to avoid computational problems). Quads are below the lower loop, with $\varphi \lesssim 2$. Doubles are between the two loops, with $2 \lesssim \varphi \lesssim 6$.[2] The values are model-dependent—for example, a steeper model will have both loops somewhat higher. Also, the value of $(\theta_1 + \theta_2)^2$ depends on the source position: smaller for sources along the long axis of the potential, larger for sources perpendicular to that axis. But with these qualifications, Figure 4 shows that the general ranges of φ, including the separation of quads and doubles, is just as it is in the data, and there is no evidence that the observed systems come from drastically different populations of lenses.

---

[2] Note that Figure 4 does not show a probability distribution, unlike related plots in Oguri et al. (2002). The aim in Figure 4 is simply to show the separation of φ for quads and doubles.
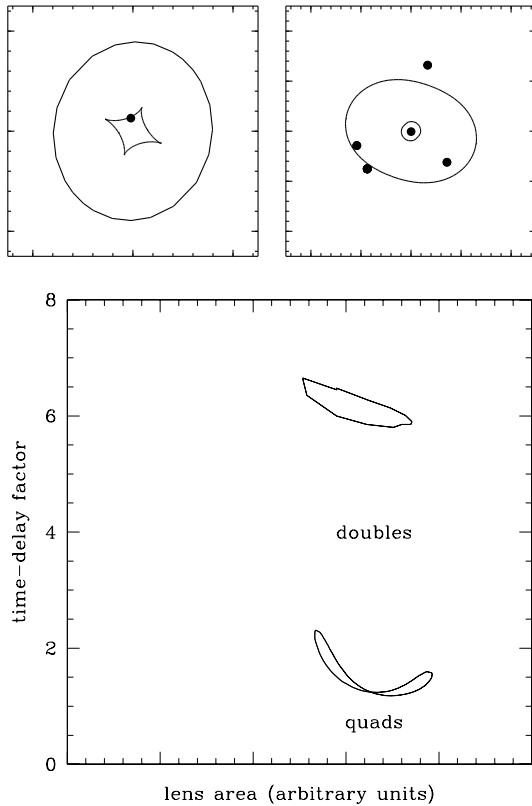
Fig. 4. Computation of φ values from a simple model of 1115+080, taken from Saha & Williams (2003). The top two panels show an image morphology similar to 1115+080, and the corresponding source position. The lower panel shows φ against $(\theta_1 + \theta_2)^2$ for source positions along the two caustics. [The horizontal axis is not labelled because $(\theta_1 + \theta_2)^2$ has arbitrary units: arcsec^2, steradians, etc.] The lower loop corresponds to the diamond caustic and the upper loop corresponds to the outer caustic. Hence quads are below the lower loop and doubles are between the two loops.

5. Summary

We see in this paper a new interpretation of lensing time delays: Δt is H_0^{-1} shrunk by the lens's covering factor on the sky, times a number of the order of unity. On separating off a redshift dependent-term (also of order unity) we are left with a number φ (say) that summarizes the dependence on details of the lens and lens configuration.

Using these ideas, we can rescale the observed time delays for the nine currently-measured systems. The observed time delays range over a factor of 40, but the rescaled delays range over a factor of 5. The latter is the inferred range of φ, and moreover it appears that $\varphi \lesssim 2$ for quads and $2 \lesssim \varphi \lesssim 6$. Reassuringly, the same spread in φ is reproduced by a simple model.

Using rescaled time-delays we can also test the hypothesis that the observed lenses all belong to a generalized-isothermal family. This hypothesis is ruled out: it over-predicts time delays for large lenses. On the other hand, there is no indication that the known time-delay systems come from drastically different types of lenses.

Appendix A: Significance of trends

In Figures 1 to 3 we have some points (x_i, y_i) and we want to know whether there is any trend in the scatter. There are many statistical tests relating to the significance of trends in data, but none of the standard ones address quite this question. However, it is not difficult to design a suitable statistical test. Let us pose the question: what is the probability of improving the fit to $y =$ constant by shuffling the y_i? If nearly all shufflings reduce the |slope| we would conclude that the data have a trend.

In the familiar straight-line fit, the slope is monotonic in $\sum_i x_i y_i$. Hence as a statistic, $\sum_i x_i y_i$ is equivalent to the slope.

In the main text, I use the phrase "significant at the 95% level" to mean that 5% of shufflings increase the |slope|. Statisticians might use a phrase like "p-value of 95%".

References

Barkana, R. 1997, ApJ, 489, 21
Berstein, G., & Fischer, P. 1999, AJ, 118, 14
Biggs, A. D., Browne, I. W. A., Helbig, et al. 1999, MNRAS, 304, 339
Burud, I., Hjorth, J., Jaunsen, A. O., et al. 2000, ApJ, 544, 117
Burud, I., Courbin, F., Magain, P., et al. 2002, A&A, 383, 71
Burud, I., Hjorth, J., Courbin, F., et al., 2002, A&A, 391, 481
Cohen, A. S., Hewitt, J. N., Moore, C. B. & Haarsma, D. B., 2000 ApJ, 545, 578
Courbin, F., Saha, P. & Schechter, P. L. 2003, in Lecture Notes in Physics, 608, Gravitational Lensing: An Astrophysical Tool, ed. F. Courbin & D. Minniti (Berlin: Springer), 1
Fassnacht, C. D., Xanthopoulos, E., Koopmans L. V. E. & Rusin D. 2002 ApJ, 581, 823
Hjorth, J., Burud, I., Jaunsen, A. O., et al. 2002, ApJ, 572, L11
Keeton, C. R.; Falco, E. E.; Impey, C. D., et al. 2000, ApJ, 542, 74
Kochanek, C. S., Falco, E. E., Impey, C., et al. 1998, cfa-www.harvard.edu/glensdata
Kochanek, C. S. 2003, ApJ, 583, 49
Kundić, T., Turner, E. L., Colley, W. N., et al. 1997, ApJ, 482, 75
Lovell, J. E. J., Jauncey, D. L., Reynolds, et al. 1998, ApJ, 508, L51
Nityananda, R. 1990, Current Science 59, 1044
Oguri, M., Taruya, A., Suto, Y. & Turner, E. L., 2002, ApJ, 568, 488
Oscoz, A., Alcalde, D., Serra-Ricart, M., et al. 2001, ApJ, 552, 81
Raychaudhury, S., Saha, P. & Williams, L. L. R. 2003, AJ, 126, 29
Refsdal, S. 1964, MNRAS, 128, 307
Saha, P. & Williams, L. L. R. 2003, AJ, 125, 2769
Schechter, P. L., Bailyn, C. D., Barr, R., et al. 1997, ApJ, 475, L85
Witt, H. J., Mao, S. & Keeton C. R. 2000, ApJ, 544, 98